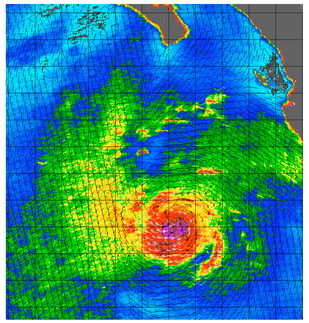
National Aeronautics and Space Administration

Earth Science Datacasting: Informed Pull and Visualization

Andrew Bingham, Timothy Stough, Bob Deen, Kevin Hussey, Sean McCleese
Jet Propulsion Laboratory, California Institute of Technology

Contact
Andrew.Bingham@jpl.nasa.gov
stough@jpl.nasa.gov

Introduction



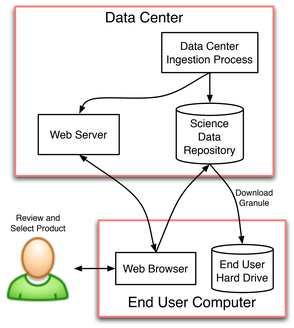
The ability to automatically and continuously download only data that meets a predefined need and instantaneously visualize it on a local computer is a concept that has yet to be realized. Our solution of earth science datacasting is based on the popular concept of podcasting. We give scientists the capability to download only content that meets their particular requirements.

Our notion of earth science datacasting is not just a data pull mechanism, but also a visualization capability. Users will have the ability to subscribe to RSS-based metadata feeds, automatically select, download and manage data that meets their need, and then visualize the data in context with other data sets.

Operational Scenario

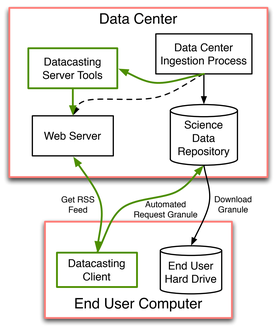
Today

Scientists must navigate their web browser to a data center, review the center's offerings, and manually download the granules of interest. As new granules are made available, the scientist must return to the site again and again to get updated data.

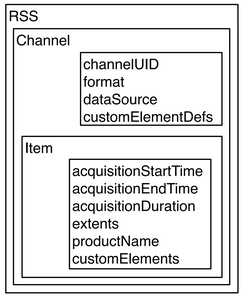


Our Innovation

- Announce new data with an extended RSS feed
- Filter and automatically download only relevant data
- Track multiple sources
- Aggregate and visualize heterogeneous data



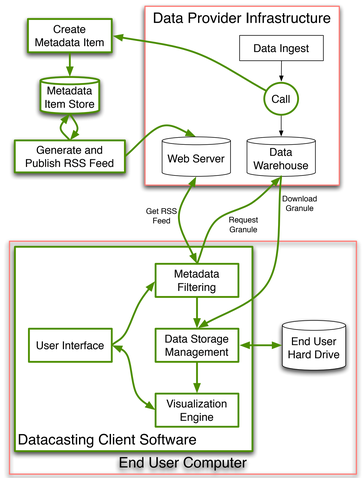
System Architecture



Datacasting Extensions to RSS

RSS, Really Simple Syndication, is an XML based format for distributing content and metadata over the internet. It is commonly used to share headlines, or, in the case of Podcasting, audio or video.

We extend RSS to add earth science specific metadata which can be filtered to select relevant data. Our extensions increase functionality, but do not reduce its compatibility with other RSS feed readers.

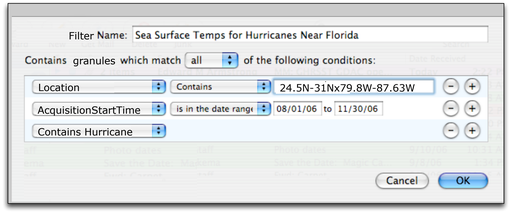


An RSS feed/channel is updated with items containing metadata for newly available granules. The client software subscribes to the feed, selects items based on metadata filters, automatically downloads the associated granule, and stores it for later analysis. Using format information specified in the channel (via ESML), stored data can be loaded, visualized, and overlaid by the client.

Client

- **Powerful filtering selects only desired data**
- **Aggregation of multiple data streams**
- **Reads and understands many data formats**
- **Visualizes and overlays different data types with fully integrated/included tools**
- **Familiar email application-like interface**

A scientist may, for example, be interested in Sea Surface Temperatures associated with hurricanes near the Florida coast. They would use our application to filter an SST data stream and automatically download relevant granules.

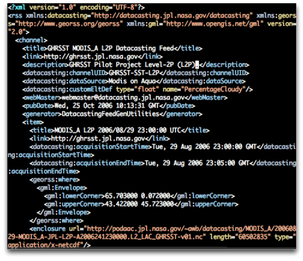


The filter creation dialog above, inspired by the Mac Mail application, allows the user to easily specify complex sets of conditions to select data granules for downloading.

Server

- **Plug and play installation at data centers**
- **Automatically publish availability of new data**
- **Fully RSS 2.0 compatible feed**
- **Uses existing web server configurations**

The datacasting server ingests metadata for each new data granule. The metadata items are queued and the RSS feed is generated regularly.



RSS feeds typically act like conveyer belts, dropping old items as new ones are added. A queuing policy controls which new items are included in the feed and which are removed. All items are archived after removal from the queue.

The RSS feed is simply an XML document which is distributed on the data center's existing web server.

Deployments/Partnerships

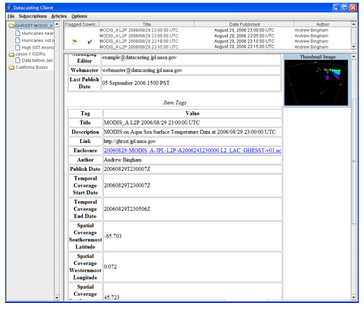
We are working with the GDAC's Global High Resolution Sea Surface Temperature (GHRSST) Project to set up a datacast of their level 2 products and ancillary data.

We are also partnering with the Physical Oceanography DAAC to datacast products formed by merging the NOAA active hurricane list with available physical oceanography data products.

Progress

Our project has been developing the datacasting concept since April, 2006. We have made the following progress:

- Drafted system requirements and architecture
- Developed and documented our extensions to RSS
- Developed and deployed a beta version of the server
  - Implemented metadata item ingestion and storage
  - All datacasting RSS elements available except *format*
  - Feed generation of validated RSS
  - Simple queuing policy for items based on time or number of items published



- Began development of the client application
  - 3-pane design showing feeds, articles, and an article detail view
  - Parses and displays datacasting and other RSS feeds
  - Utilizes datacasting specific elements in feed
  - Visualizes preview image if included in feed

The work presented in this poster is funded by the ACCESS program (SMD ROSES - NNH05ZDA001N A.26) and is being carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA